Journal of Risk and Uncertainty, 17:87–119 (1998)

Dynamic Choice and NonExpected Utility

RAKESH SARIN
The Anderson Graduate School of Management, UCLA, Los Angeles, CA

PETER P. WAKKER
CentER for Economic Research, Tilburg University, The Netherlands

Abstract

This paper explores how some widely studied classes of nonexpected utility models could be used in dynamic choice situations. A new "sequential consistency" condition is introduced for single-stage and multi-stage decision problems. Sequential consistency requires that if a decision maker has committed to a family of models (e.g., the multiple priors family, the rank-dependent family, or the betweenness family) then he use the same family throughout. Conditions are presented under which dynamic consistency, consequentialism, and sequential consistency can be simultaneously preserved for a nonexpected utility maximizer. An important class of applications concerns cases where the exact sequence of decisions and events, and thus the dynamic structure of the decision problem, is relevant to the decision maker. It is shown that for the multiple priors model, dynamic consistency, consequentialism, and sequential consistency can all be preserved. The result removes the argument that nonexpected utility models cannot be consistently used in dynamic choice situations. Rank-dependent and betweenness models can only be used in a restrictive manner, where deviation from expected utility is allowed in at most one stage.

Key words: Dynamic choice, nonexpected utility, dynamic consistency, consequentialism

JEL Classification: D81

In the past two decades, several nonexpected utility (nonEU) models have been presented in the economic literature. These models permit empirically observed preference patterns that are otherwise inconsistent with the classical expected utility theory. This paper explores how such nonEU models could be employed in dynamic choice situations. The main focus of this paper is with decisions under uncertainty where probabilities are not given. Throughout we will use the common term "expected utility" for both the cases of risk (probabilities are given) and uncertainty.

An example of a dynamic choice situation will be discussed in Section 1. In the first stage a production decision is to be made. In the second stage a marketing decision will be made that is contingent on whether the unit production cost turns out to be high or low. In dynamic choice, some decisions (marketing) are made only after the resolution of an uncertainty (unit cost high or low). We can therefore compare the consistency of an initial plan or decision strategy with the actual choices at subsequent decision nodes. Consistency in dynamic choice situations has been a significant argument in favor of expected utility models. This is because an expected utility maximizer's planned choice at the beginning will coincide with the choices made upon arriving at intermediate decision nodes. So an expected utility maximizer does what he plans and is therefore dynamically consistent.

Whether consistency can be attained with nonEU models has been a topic of much debate and controversy in recent years. It is clear that a naive implementation of nonEU models results in inconsistent dynamic behavior (Hammond 1988). In response to this potential criticism of nonEU models, several authors have suggested ways to implement such models in dynamic choice situations. We discuss these suggestions after presenting our results.

In our results, consistency principles need only be imposed on a particular decision tree at hand. The principles may, but need not, hold in other trees involving other (orderings of) events. This flexibility increases the applicability and mathematical generality of our results. For example, the exact sequence of decisions and events can, but need not, be relevant to the decision maker. Our results can be applied if dynamic consistency and consequentialism are universally satisfied (as in Chew & Epstein 1989, Segal 1990, Luce & Fishburn 1991, Grant, Kajii, and Polak 1998). They can also be applied if these two conditions are not universally satisfied, but only in the decision tree under consideration. The latter case is illustrated by the example of Section 1 (compare Figure 5 and endnote 1 in Section 8).

Folding back (other terms are backward induction or dynamic programming) is an attractive procedure to evaluate dynamic decisions. The method recursively substitutes certainty equivalents at chance nodes and optimal certainty equivalents at decision nodes, starting at terminal nodes. (Substitution of certainty equivalents is one way of formulating folding back, and is the formulation used throughout this paper.) Folding back is assumed in the majority of works in the economics of temporal preference (Strotz 1956, Kreps and Porteus 1979, Johnsen and Donaldson 1985, Chew and Epstein 1989, Epstein and Zin 1989, Karni and Safra 1990). Therefore we study the possibility of using folding back with nonEU in specific decision situations

The new condition in this paper is sequential consistency, which ensures consistency in application of a family of models to alternative representations of the same decision problem. Sequential consistency requires that if a decision maker has committed to a family of models (e.g., the rank-dependent family), then he should use that family of models throughout. That is, he should use the same family of models both in an evaluation of a two-stage tree using folding back and in a direct evaluation of associated strategies in a single-stage tree. Sequential consistency is similar to the meta-principle employed by Epstein and Le Breton (1993). Their condition requires a richer domain of decision trees but does not require full-force folding back. A detailed discussion of related conditions can be found in Section 8.

We obtain the following implications of sequential consistency and folding back in a given decision tree. For one important class of nonEU models (multiple priors), these conditions can be preserved in a same manner as for expected utility. Rank-dependent and betweenness models can only be used in a restrictive manner, where in at most one stage deviation from expected utility is allowed.

In Section 1 we present an example of a dynamic choice problem and discuss sequential consistency in the context of the example. We also discuss the relevant notions such as dynamic consistency and consequentialism. In Sections 2, 3, 4, 5, and 6, results are reported for the family of multiple priors models, rank-dependent models, the family of betweenness models, the family of weighted utility models, and the general family of transitive preferences. For simplicity, the results have been presented for two stage decision trees with two-branch event nodes. Section 7 demonstrates that the results immediately extend to multiple stages and events. Section 8 provides a discussion and a comparison of our results with the literature. Conclusions are presented in Section 9. All proofs are contained in the appendix. The appendix also demonstrates that our theorems can be reinterpreted as a study of separability in static nonEU models (e.g., Corollary B.3).

1. The dynamic decision problem

An example of a dynamic choice problem is given in Figure 1a. A production decision is made and the unit cost is observed. The unit cost could be high (E) or low (E^c). Next a marketing decision is made that is contingent on the observed unit cost. The marketing decision involves the uncertainty that the market size could be big or small. We write s_1 for (high unit production cost, big market size), s_2 for (high unit production cost, small market size), s_3 for (low unit production cost, big market size), and s_4 for (low unit production cost, small market size).

Suppose that in Figure 1a, at decision node 1, one may choose f which takes one to decision node 3 if E occurs and to decision node 5 if E^c occurs. At decision node 3 one may choose g which yields $\$x_1$ if s_1 (big market size) and $\$x_2$ if s_2 (small market size). Figure 1a depicts the dynamic choice situation where at nodes 3 and 5 choices are made after resolution of the uncertainty concerning the unit cost.

The notation $s_1, \ldots, s_4$ is chosen here for consistency with the formal analysis presented later in the paper, where $\{s_1, \ldots, s_4\}$ is a state space. Formally, the partitioning $\{s_1,s_2\}$ of E also applies if c is chosen at node 3, although it then is of no interest to the decision maker which of s_1, s_2 is true. Similarly, the partitioning $\{s_3,s_4\}$ of E^c also applies if d is chosen at node 5 and similar partitionings apply if k is chosen at node 1.

We also study the *single-stage choice situation* that results from Figure 1a if the decision maker must commit beforehand, before the uncertainty about the production cost gets resolved, to choices at nodes 3 and 5. Then the decision maker chooses, beforehand, between strategies. A strategy may be viewed as a set of instructions that can be given to an agent and that unambiguously tell him what to do at each decision node. An example is (choose f, if E choose g, if E^c choose h), abbreviated (f,g,h). The single-stage choice situation is illustrated in Figure 1b. Figure 1b should be interpreted as preserving the information on the dynamic nature (exact sequence of events) of the choice situation, including the timing of the resolution of uncertainty. The figure should therefore be distinguished from a normal form tree in which the exact sequence of events does not matter.

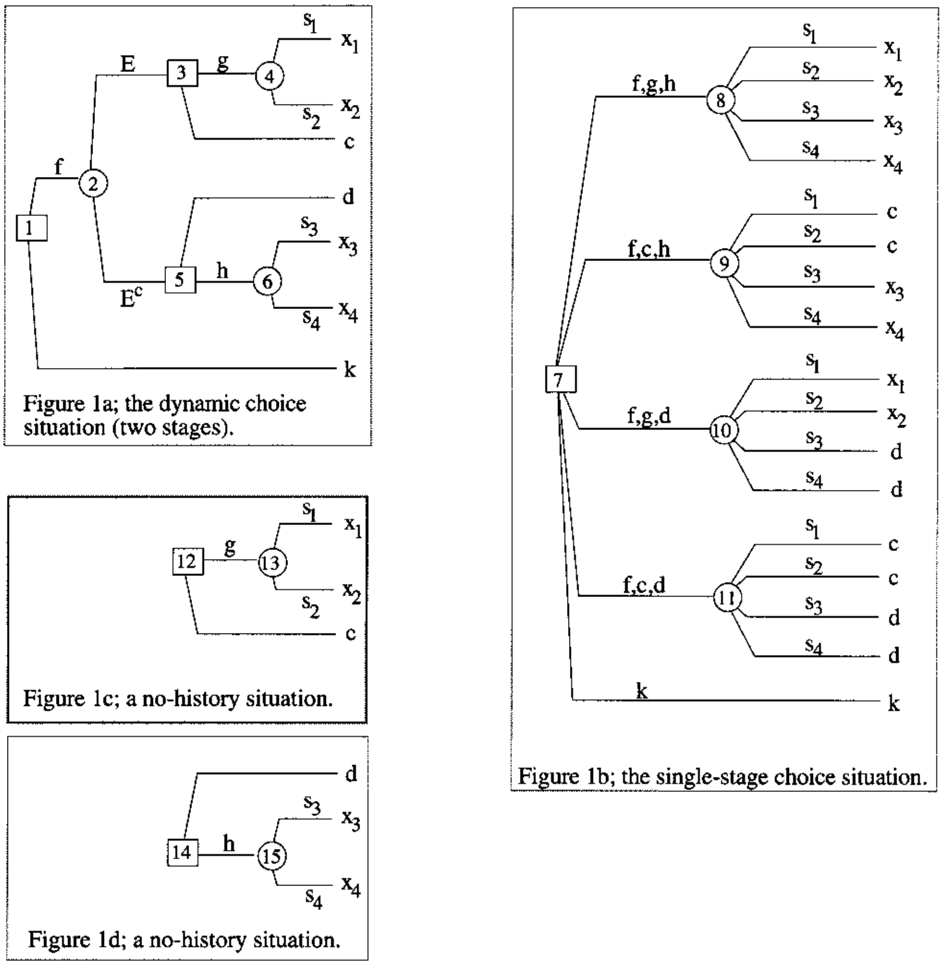

Figure 1a; the dynamic choice situation (two stages).

Figure 1b; the single-stage choice situation.

Figure 1c; a no-history situation.

Figure 1d; a no-history situation.

Figure 1. An example of a dynamic choice problem. *Dynamic consistency*: the choices in Figure 1a agree with the choice in Figure 1b. *Consequentialism*: the choices in Figure 1a agree with the choices in Figures 1c and 1d.

The requirement that the choices in the dynamic (Figure 1a) and the single-stage (Figure 1b) situation agree, is called *dynamic consistency* (Epstein 1992, Figure 1.6). So, if the strategy (f,g,d) is preferred at node 7, then f is chosen at node 1, g is chosen at node 3, and d is chosen at node 5. Similarly, if g is preferred at node 3, then strategies (f,c,h) and (f,c,d) are not chosen at node 7. For alternative models that permit changing preferences over time or allow for flexibility in future option see Strotz (1956) and Kreps and Porteus (1979).

Another condition for dynamic choices is *consequentialism*. It requires that a choice at a decision node, such as node 3, is made independently of risks forgone in the past. So a choice at node 3 agrees with that at node 12 and a choice at node 5 agrees with that at node 14. This means that a choice at node 3 is independent of the values x_3 and x_4, and a choice at node 5 is independent of the values x_1 and x_2.

Summarizing, dynamic consistency requires agreement of choices in Figures 1a and 1b, and consequentialism requires agreement of choices in Figures 1a and 1c, 1d.

If dynamic consistency and consequentialism are satisfied, then dynamic decisions can be determined through the folding back procedure. An application of this procedure to the decision tree in Figure 1a will require transporting oneself to chance node 4 and substituting there a certainty equivalent; next at decision node 3 the highest available certainty equivalent (the other being c) is chosen and substituted for this decision node. Similarly, working backwards from the certainty equivalent at node 6, the certainty equivalent for node 5 is computed (the higher of d and the certainty equivalent at node 6). At node 2, the certainty equivalent of the lottery is substituted which yields the certainty equivalent computed at node 3 if E occurs and the certainty equivalent computed at node 5 if E^c occurs. Finally, at node 1 a choice is made between k and the certainty equivalent at node 2.

For the evaluation of decision trees, we propose a new condition, sequential consistency. It is a meta-principle for the folding back procedure. Suppose the decision maker commits himself to using a family of models *M*, say the rank-dependent family. Thus he uses the family *M* to evaluate certainty equivalents in the folding back procedure. The meta-principle requires that he use this family *M* to evaluate decision strategies in the single-stage tree as well. In other words, it requires that the evaluation of a strategy obtained by using a family *M* through the folding back procedure should coincide (in the sense of giving the same certainty equivalents) with a direct evaluation of the same strategy in the single-stage tree using the same family *M*. This means that, if at nodes 6, 4, and 2 in Figure 1a, certainty equivalents are obtained through a form from *M*, then similar things hold at nodes 8, 9, 10, and 11.

Sequential consistency is normatively appealing once folding back is assumed. If a decision maker accepts the axioms of a model *M* then, while he might use different members of the family at different stages, the preference functionals should always be from *M*. For example, assume the decision maker considers comonotonic independence and the other axioms of rank-dependent models as normatively convincing. He therefore uses rank-dependent forms to derive certainty equivalents at nodes 6, 4, and 2. It would then be undesirable if violations of comonotonic independence or the other rank-dependent axioms were generated in Figure 1b.

A pragmatic reason for accepting sequential consistency can be that a functional form dictated by *M* is often used to derive results and explain economic behavior such as risk aversion and portfolio preference. It is convenient if these results and explanations can be applied at all decision nodes regardless of how we formulate the decision problem.

For the presentation of our formal results, we focus on the evaluation of strategy (f,g,h) in Figure 1. This strategy has been depicted in Figure 2, with Figure 2a corresponding to Figure 1a and Figure 2b corresponding to Figure 1b. This strategy (f,g,h) is also denoted

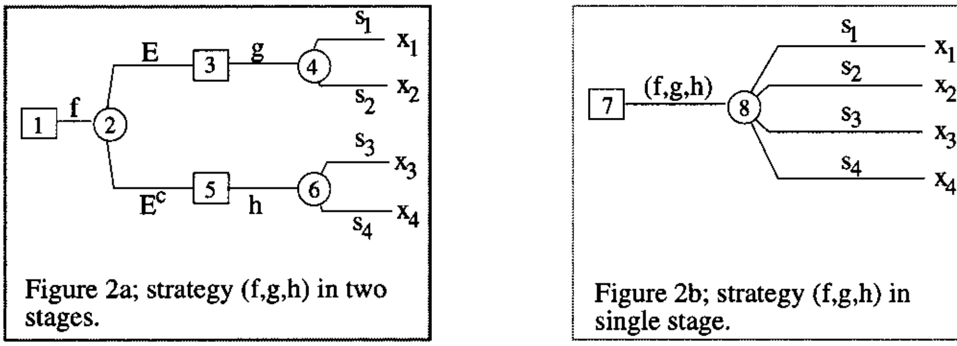

Figure 2a; strategy (f,g,h) in two stages.

Figure 2b; strategy (f,g,h) in single stage.

Figure 2. Depicting one strategy.

as $(x_1,\cdots,x_4)$, which is taken as a generic notation for strategies. The *outcome space* is denoted by $\mathscr{C}$, where $\mathscr{C}$ is an interval. The set of strategies is $\mathscr{C}^4$. By $\succcurlyeq$ we donate preferences over strategies. So for the domain of $\succcurlyeq$, $s_1,\cdots,s_4$ are fixed and the outcomes $x_1,\cdots,x_4$ vary over $\mathscr{C}$. The term "event tree" could be used to express that in our domain only the events in the tree are kept fixed but the outcomes are still varied. Trees are to be interpreted in that sense in this paper. *Strong monotonicity* means that increasing any outcome of a strategy leads to a strictly preferred strategy. It implies that there are no null states. The following assumption is made throughout the paper (except in Appendix F).

Assumption 1.1. The preference relation $\succcurlyeq$ satisfies transitivity, completeness, continuity, and strong monotonicity. □

The assumption implies strictly increasing and continuous utility functions in all models throughout this paper. We assume that the person has decided that evaluations of choice alternatives should be described by a specific class of nonexpected utility (nonEU) models, denoted *M*. For instance, we consider hereafter the cases where *M* is the class of the multiple priors forms, the rank-dependent forms, or the betweenness forms. *M* should be rich enough to contain evaluation forms for all strategies $(x_1,\cdots,x_4)$ and also for all relevant substrategies such as pairs (x_1,x_2) and (x_3,x_4). We will not pursue a formalization of the family *M*, which would require tools from logic. In the theorems in this paper, *M* and its domains will be specified, at which stage the analysis becomes fully formalized.

It turns out to be most convenient for our analysis if we let the evaluating forms in *M* be expressed in terms of certainty equivalents. For instance, if *M* is the expected utility (EU) family, then a form from *M* for evaluating strategies $(x_1,\cdots,x_4)$ adopts probabilities $p_1,\cdots,p_4$ for $s_1,\cdots,s_4$ and a utility U, and then evaluates the strategy $(x_1,\cdots,x_4)$ by its certainty equivalent $U^{-1}(\sum_{j=1}^{4} p_jU(x_j))$. A form from that same family *M* for evaluating substrategies (x_1,x_2) at node 4 in Figure 1a uses probabilities q_1,q_2 and a utility $\tilde{U}$, and assigns to (x_1,x_2) the certainty equivalent $\tilde{U}^{-1}(q_1\tilde{U}(x_1) + q_2\tilde{U}(x_2))$. In general, $\tilde{U}$ may be different than U, but in our main results $\tilde{U}$ will always coincide with U, due to the assumptions made there. In agreement with Assumption 1.1, the models in *M* are assumed to satisfy continuity and monotonicity throughout this paper.

We assume that strategies are evaluated by means of a folding back procedure in Figure 2a, as follows. At node 4 the decision maker replaces (x_1,x_2) by a certainty equivalent $V_E(x_1,x_2)$, at node 6 he replaces (x_3,x_4) by a certainty equivalent $V_{E^c}(x_3,x_4)$, next these values are substituted at nodes 3, 5, and, finally, at nodes 2 and 1 he replaces the pair $(V_E(x_1,x_2),V_{E^c}(x_3,x_4))$ by a certainty equivalent $V(V_E(x_1,x_2),V_{E^c}(x_3,x_4))$. Further, V_E, V_{E^c}, and V are forms from *M*. The procedure just described, *folding back with respect to the family M*, is displayed in (1.1).

$$\textit{Strategies } (x_1,\cdots,x_4) \textit{ are evaluated by } V(V_E(x_1,x_2),\ V_{E^c}(x_3,x_4)),\quad \textit{where } V_E,\ V_{E^c}, \textit{ and } V \textit{ are forms from } M. \tag{1.1}$$

We write $W(x_1,\cdots,x_4)$ for the certainty equivalent of strategy $(x_1,\cdots,x_4)$ in Figure 2b. Because of consequentialism and dynamic consistency, the procedure for evaluating strategies as in (1.1) provides certainty equivalents of strategies both in Figures 2a and 2b, and therefore $V(V_E(x_1,x_2),V_{E^c}(x_3,x_4)) = W(x_1,\cdots,x_4)$. Sequential consistency requires that this functional, when considered a functional on quadruples, should be an element of *M*. In other words, *sequential consistency* requires that (1.1) implies:

$$W(x_1,x_2,x_3,x_4) \textit{ is a form from } M. \tag{1.2}$$

For example, if *M* is the rank-dependent family, then sequential consistency requires that W satisfies comonotonic independence and the other axioms of the rank-dependent family. We call (1.2) the *single-stage evaluation* approach (with respect to *M*).

EU preferences satisfy (1.1) and (1.2) with respect to the set *M* of EU forms. For instance, if $W(x_1,x_2,x_3x_4) = U^{-1}(\sum_{j=1}^{4} p_jU(x_j))$, then $V_E(x_1,x_2) = U^{-1}(q_1U(x_1) + q_2U(x_2)) = y$ for conditional probabilities $q_1 = p_1/(p_1 + p_2)$, $q_2 = p_2/(p_1 + p_2)$, $V_{E^c}(x_3,x_4) = U^{-1}(q_3U(x_3) + q_4U(x_4)) = z$ for conditional probabilities $q_3 = p_3/(p_3 + p_4)$, $q_4 = p_4/(p_3 + p_4)$, and $V(y,z) = U^{-1}((p_1 + p_2)U(y) + (p_3 + p_4)U(z))$.

Let us emphasize that the dynamic information of the decision situation is preserved in the single-stage evaluation and that therefore W can depend on the order of the resolution of uncertainty. If that order were changed (e.g., as in Figure 5 hereafter), then another function W* might possibly be adopted to evaluate quadruples $(x_1,\ldots,x_4)$. In other words, we do not assume normal-form equivalence.

Dynamic consistency and consequentialism, thus the folding back procedure, imply for the single-stage choices in Figure 1b that the events E $(=\{s_1,s_2\})$ and E^c $(= \{s_3,s_4\})$ are separable. *Separability* of $\{s_1,s_2\}$, i.e., event E, means that $W(x_1,\cdots,x_4)$ can be written as $F(V_E(x_1,x_2),x_3,x_4)$, where F is strictly increasing in $V_E(x_1,x_2)$. Separability of other events is defined similarly. Separability of preferences means separability of all events. Separability of E and E^c with respect to the single-stage choices is the only implication of the dynamic choice conditions (consequentialism and dynamic consistency) used in our mathematical theorems. Therefore our results could be reinterpreted as an analysis of separability for single-stage nonEU models. For rank-dependence, further comments are provided at the end of Appendix B.

In the next sections we explore whether (1.1) and (1.2) can be simultaneously satisfied for other classes *M* of models for decision under uncertainty and risk than expected utility.

2. Multiple priors

This section shows that the multiple priors family can satisfy sequential consistency without reducing to expected utility at any stage. Decision under uncertainty assumes a *state space* Ω and an outcome space $\mathcal{C}$. Ω is assumed to be $\{s_1,\ldots,s_4\}$ up to Section 7. *Acts* f map states to outcomes and the preference relation $\succcurlyeq$ is over acts. The *multiple priors model* holds if there exists a set $\mathcal{P}$ of probability measures on the state space and a *utility* $U : \mathcal{C} \to \mathrm{IR}$, such that

$$f \mapsto min_{P\in\mathcal{P}} \int_{\Omega} U(f(\omega))dP$$

represents $\succcurlyeq$. The family *M* of multiple priors forms contains all such representations for all state spaces and outcome spaces. The model has been used by Wald (1950). A characterization in terms of preference conditions has been provided by Gilboa and Schmeidler (1989). The model is compared to Choquet expected utility by Klibanoff (1995) and Ben-Porath, Gilboa, and Schmeidler (1997).

Now consider the decision tree in Figure 1a. Assume that for node 2 a family $\mathcal{P}_2$ of probability measures over $\{s_1,\ldots,s_4\}$ is given, for node 4 a family $\mathcal{P}_4$ of probability measures over $\{s_1,s_2\}$, and for node 6 a family $\mathcal{P}_6$ of probability measures over $\{s_3,s_4\}$. Assume that the decision maker does folding back according to the multiple priors model. Consider the model in Figure 1b. In determining the family of prior probabilities at node 7, we will assume probability multiplication. Hence the family $\mathcal{P}_7$ at node 7 consists of all probability distributions that result from taking one element of $\mathcal{P}_2$, one element of $\mathcal{P}_4$, and one element of $\mathcal{P}_6$, and then using probability multiplication. The family $\mathcal{P}_7$ is called the *reduced family* of probability measures. It suggests a stochastic independence of the separate chance nodes in Figure 1a, where the probability measure at any chance node does not affect the probability measures at the other nodes. That allows for the preferential separability needed for sequential consistency.

Theorem 2.1. Suppose that Assumption 1.1 holds, that folding back (Formula 1.1) holds with respect to the family *M* of multiple priors forms, and that the same utility function is used at nodes 2, 4, and 6. Then sequential consistency holds with respect to *M* if the multiple priors form W in (1.2) uses the reduced family of probability measures and the same utility function. □

3. Rank-dependent utility

In rank-dependent utility (RDU), nonadditive or nonlinear measures are used. For decision under uncertainty, the model considered here, RDU is usually called "Choquet expected utility." The state space Ω, the outcome space $\mathscr{C}$, and acts are as before. A *capacity* v is a function on the subsets of Ω that satisfies $v(\varnothing) = 0$, $v(\Omega) = 1$, and that is *monotonic* with respect to set inclusion, i.e., $A \supset B \Rightarrow v(A) \geq v(B)$. *Choquet expected utility* (*CEU*) holds if a capacity v on Ω and a *utility* $U : \mathscr{C} \to \mathrm{IR}$ exist such that

$$f \mapsto \int_{IR^+} v\{\omega \in \Omega: U(f(\omega)) \geq \tau\} d\tau + \int_{IR^-} [v\{\omega \in \Omega: U(f(\omega)) \geq \tau\} - 1] d\tau$$

represents $\succcurlyeq$. The integral is the *Choquet expected utility* (*CEU*) of f and is also denoted as $\int_\Omega U^\circ f dv$. If v is additive, then CEU reduces to the classical expected utility (EU), as one verifies by partial integration. For the state space $\{s_1,\cdots,s_4\}$, as considered in the single-stage evaluation, CEU of an act ("strategy") x with $x_1 \geq \cdots \geq x_4$ can be rewritten as the sum

$$\sum_{j=1}^{4} \pi_j U(x_j), \tag{3.1}$$

where the *decision weights* π_j are defined as $v(\{s_1,\cdots,s_j\}) - v(\{s_1,\cdots,s_{j-1}\})$. For the state space $\{s_1,s_2\}$, considered in the folding back procedure at node 4 of Figure 2a, CEU of an act ("subact") (x_1,x_2) with $x_1 \geq x_2$ is $\tilde{v}(\{s_1\})\tilde{U}(x_1) + (1 - (\{s_1\}))\tilde{U}(x_2)$, where $\tilde{v},\tilde{U}$ may be different from v,U as used in the single-stage evaluation (3.1).

The *family M of rank-dependent forms*, i.e., the *RDU* family, contains all CEU representations for all state spaces and outcome spaces. A special case of decision under uncertainty is decision under risk, where probabilities are given on Ω so that acts can be identified with the probability distributions over outcomes (Wakker 1990). Thus, the RDU family includes the models for decision under risk introduced by Quiggin (1981).

A characterization in terms of preference conditions of CEU with continuous utility, the model considered here, has been provided by Wakker (1989). Now we present the implication of sequential consistency for folding back with the RDU family.

Theorem 3.1. Suppose that Assumption 1.1 holds and that folding back (Formula 1.1) holds with respect to the family *M* of rank-dependent forms. Then the following three statements are equivalent:

(i) Sequential consistency holds with respect to *M*.
(ii) V is an expected utility form and V, V_E, and V_{E^c} use the same utility.
(iii) There exist a utility $U : \mathscr{C} \to \mathrm{IR}$, probabilities P(E) and $P(E^c)$ ($P(E) + P(E^c) = 1$), and capacities v_E on $\{s_1,s_2\}$ and v_{E^c} on $\{s_3,s_4\}$, such that $\succcurlyeq$ can be represented by

$$(x_1,\cdots,x_4) \mapsto P(E)CEU_E(x_1,x_2) + P(E^c)CEU_{E^c}(x_3,x_4),$$

where CEU_E denotes Choquet expected utility with respect to U and ν_E, and CEU_{E^c} denotes Choquet expected utility with respect to U and ν_{E^c}.

□

We briefly describe the proof, elaborated in the Appendix, of the most interesting implication, i.e., that (i) implies that V is an EU form. The key step is to show that the decision weights of s_1 and s_2 do not depend on the outcomes x_3 and x_4. To see this, note that an indifference $(x_1,x_2,x_3,x_4) \sim (x_1 + \epsilon,x_2 - \delta,x_3,x_4)$, for $x_1 + \epsilon > x_1 > x_2 > x_2 - \delta > x_3 > x_4$ should, be separability of $\{s_1,s_2\}$, be kept if x_3 is replaced by x_3' such that $x_1 > x_3' > x_2$. In the CEU formula (3.1), this replacement does not affect the decision weight of state s_1. Therefore, to preserve the indifference, the decision weight of s_2 should neither be affected by the replacement. By similar reasonings for other cases it follows that the decision weights of s_1 and s_2 are indeed independent of the outcomes x_3 and x_4. We define $P(E) = \nu(E) = 1 - P(E^c), \nu_E = \nu/P(E), \nu_{E^c} = \nu/P(E^c)$, and the representation in (iii) follows from substitution.

We display the most important implication of Theorem 3.1:

Corollary 3.2. If folding back and sequential consistency hold with respect to the family of rank-dependent forms, then the decision maker is free to use a nonadditive capacity in the second stage but must maximize expected utility in the first stage. □

The corollary may be useful when first stage events are ("roulette wheel") events that do not involve ambiguity and those in the second stage are ("horse race") events that do involve ambiguity. Consistent with other authors, e.g. Schmeidler (1989), a person may maximize EU with respect to unambiguous events and deviate from EU for horse race events. Schmeidler's paper, which introduced CEU, assumes that the unambiguous events are in the second stage and the horse-race events are in the first stage. Our result suggests a reversal of the stages to maintain sequential consistency (compare Corollary F.2 in Appendix F). Eichberger and Kelsey (1996) assumed dynamic consistency but not consequentialism and considered choices conditional on an event E by fixing outcomes given E^c. Under some nontriviality assumptions, they showed that CEU with "strict uncertainty aversion" in the single-stage situation excludes CEU for the choices conditional on E. Following Schmeidler, they assumed that outcomes are probability distributions over "prizes" and that with respect to these the decision maker maximizes EU. In a mathematical sense this comes down to linearity of utility in outcomes.

Finally, we give an example in which all conditions of Theorem 3.1 are satisfied but EU does not hold.

Example 3.3. Suppose that U is the identity, $P(E) = P(E^c) = 1/2$, and $\nu_E(s_1) = \nu_E(s_2) = \nu_{E^c}(s_3) = \nu_{E^c}(s_4) = 0.4$. Suppose $x_1 > \cdots > x_4$. The value of strategy $(x_1,\cdots,x_4)$ is

$$\frac{1}{2} \times 0.4 \times x_1 + \frac{1}{2} \times 0.6 \times x_2 + \frac{1}{2} \times 0.4 \times x_3 + \frac{1}{2} \times 0.6 \times x_4.$$

If, for example, $x_2 > x_1$ and $x_4 > x_3$, then the value of the strategy is

$$\frac{1}{2} \times 0.4 \times x_2 + \frac{1}{2} \times 0.6 \times x_1 + \frac{1}{2} \times 0.4 \times x_4 + \frac{1}{2} \times 0.6 \times x_3.$$

Note that the outcomes have different decision weights in the second evaluation than in the first. Under EU, the decision weights would have to be identical in the two cases. Therefore, EU does not hold in this example. □

4. Betweenness

For decision under risk the family of betweenness models has been extensively studied (Fishburn 1988). If two probability distributions are indifferent, then in the betweenness models all distributions obtained by a convex combination of these distributions are indifferent as well. Thus, in a probability simplex all indifference curves are linear but not necessarily parallel. This means that each indifference class of a betweenness model is also an indifference class of an EU model.

For decision under uncertainty no one, to our knowledge, has characterized the betweenness preferences in general. For the purpose of this paper, it is sufficient to note that betweenness holds if for each indifference class there exists an EU model such that the indifference class is also an indifference class of that EU model. For alternative indifference classes the implied EU models may differ from one another. We next give a formal definition of betweenness, extending it from decision under risk to the more general context of decision under uncertainty.

Monotonicity is assumed throughout this paper. For betweenness models, monotonicity may be problematic if outcomes are unbounded. Therefore we restrict the domain and assume that the outcome set is [0,M] for some fixed $M > 0$. Let W denote a strictly monotonic continuous function that represents $\succcurlyeq$. For a number k in the range of W, $W^{-1}(k)$ is a preference indifference class. We often use k as a superscript; k does not designate an exponent. *Betweenness* holds if for every k in the range of W, there exist probabilities $p_1^k, \cdots, p_4^k$ and a strictly increasing continuous utility U^k such that $W^{-1}(k) = \{(x_1, x_2, x_3, x_4) \in [0,M]^4 : p_1^k U^k(x_1) + \cdots + p_4^k U^k(x_4) = k'\}$ for some constant k'. The constant k' can be chosen arbitrarily by rescaling U^k. For instance, suppose we let k' be 0 for each k. Then $W(x_1, \cdots, x_4) = k$ if and only if $\sum_{j=1}^{4} p_j^k U_k(x_j) = 0$. Therefore the value k of the strategy can also be obtained as the solution in k of the implicit equation $\sum_{j=1}^{4} p_j^k U_k(x_j) = 0$. Such an implicit way of describing the betweenness family was given by Dekel (1986) and Chew (1989) for decision under risk. Folding back (1.1) with respect to the betweenness family has been discussed, in a more complex dynamic setup, by Epstein and Zin (1989, "Class 3") and Chew and Epstein (1989, Theorem 2). Epstein (1992, Section 3.4) gives a normative argument for folding back with respect to the betweenness family.

Theorem 4.1. Suppose that Assumption 1.1 holds and that folding back (Formula 1.1) holds with respect to the family M of betweenness forms. Then the following two statements are equivalent:

(i) Sequential consistency holds with respect to M.
(ii) The preference relation can be represented by a betweenness form $W(x_1,\cdots,x_4) = V(V_E(x_1,x_2),V_{E^c}(x_3,x_4))$ such that V_E and V_{E^c} are EU forms. Further, the EU models $(p_1^k,\cdots,p_4^k,U^k)$ for which $W^{-1}(k)$ is an indifference class, are such that the conditional probabilities $p_1^k/(p_1^k + p_2^k)$ and $p_2^k/(p_1^k + p_2^k)$ as well as $p_3^k/(p_3^k + p_4^k)$ and $p_4^k/(p_3^k + p_4^k)$ are independent of the indifference class parameter k. The utilities $U^k = U$ are also independent of k. V_E and V_{E^c} use the same utility U.

□

In the theorem, the decision maker maximizers EU given event E with utility U (the index k can be omitted) and conditional probabilities $p_1^k/(p_1^k + p_2^k)$ and $p_2^k/(p_1^k + p_2^k)$. Similarly, the decision maker maximizes EU given event E^c with utility U and conditional probabilities $p_3^k/(p_3^k + p_4^k)$ and $p_4^k/(p_3^k + p_4^k)$. All conditional probabilities are independent of k. The only deviation from EU is that the probability $p_1^k + p_2^k$ for event E and the probability $p_3^k + p_4^k$ for event E^c depend on the indifference class that is assigned value k.

We display the most important implication:

Corollary 4.2. If folding back and sequential consistency hold with respect to the family of betweenness forms, then the decision maker can use a betweenness model in the first stage but must maximize expected utility in the second stage. □

An example illustrates Theorem 4.1.

Example 4.3. Suppose that U is the identity function. The conditional probabilities of both s_1 and s_2, given E, are 1/2, and so are the conditional probabilities of s_3 and s_4 given E^c. Define the form V as

$$V(z_1,z_2) = \frac{2z_1 + z_2}{z_1 - z_2 + 3M}.$$

Finally, each $(x_1,\cdots,x_4)$ is evaluated by $V((x_1 + x_2)/2, (x_3 + x_4)/2))$. This form agrees with (1.1) (rewrite as

$$V(U^{-1}(\frac{U(x_1) + U(x_2)}{2}),U^{-1}(\frac{U(x_3) + U(x_4)}{2})), etc.).$$

It is not an EU form and neither is it an ordinal transform thereof. Its indifference curves $V(z_1,z_2) = k$ are straight lines of the form

$$z_2(1 + k) = (k - 2)z_1 + 3kM$$

that all pass through (−M,2M). They correspond with an indifference class of the EU model where utility is the identity and probabilities are $(1 + k)/3$ and $(2 - k)/3$ respectively. Finally, preferences are represented by

$$(x_1,\cdots x_4) \mapsto W(x_1,x_2,x_3,x_4) + \frac{2x_1 + 2x_2 + x_3 + x_4}{x_1 + x_2 - x_3 - x_4 + 6M}.$$

□

5. Weighted utility

Weighted utility is a special form of implicit utility. As implicit utility, it has been studied mainly for decision under risk (Chew 1983). Hazen (1987) studied it in a model in which the first stage deals with decision under uncertainty. In the second stage he assumed decision under risk. See also Eichberger and Grant (1997) and Lo (1996). For decision under uncertainty with a state space $\{s_1,\cdots,s_4\}$ and acts being functions from $\{s_1,\cdots,s_4\}$ to [0,M], weighted utility is defined as follows. First, the decision maker chooses subjective probabilities $p_1,\cdots,p_4$ for the states and a strictly increasing continuous utility U: [0,M] → IR, similarly to EU. Next, the decision maker chooses a "weighting function" w: [0,M] → IR_+. An act $(x_1,\cdots,x_4)$, assigning outcome x_j to state s_j, is evaluated by a convex combination of its outcome utilities $U(x_j)$, using the following formula:

$$\sum_{j=1}^{4} \frac{p_j w(x_j) U(x_j)}{\sum_{i=1}^{4} p_i w(x_i)}. \tag{5.1}$$

EU is the special case where the function w is constant. An indifference class of weighted utility is a set of acts where (5.1) is a constant k. In other words, if we replace the U (x_j)'s in the convex combination by $U(x_j) - k$, the resulting value should be zero. That does not change if we delete the denominator and thus indifference classes are described by

$$\sum_{j=1}^{4} p_j w(x_j)(U(x_j) - k) = 0.$$

This is indeed the indifference class of EU form, i.e., the one with probabilities $p_j^k = p_j$ and utility $U^k(x_j) = w(x_j)(U(x_j) - k)$. So weighted utility is an implicit utility model, where the EU models that describe indifference classes have varying utilities but all use the same probabilities. We found, however, that the folding back model in Theorem 4.1 does not permit EU descriptions of indifference classes that have varying utilities. This proves that folding back with weighted utility cannot be reconciled with the sequential consistency condition of this paper; that is, it must then reduce to EU.

Theorem 5.1. Suppose that Assumption 1.1 holds and that folding back (Formula 1.1) holds with respect to the family *M* of weighted utility forms. Then the following two statements are equivalent:

(i) Sequential consistency holds with respect to *M*.
(ii) The preference relation maximizes EU.

□

6. General transitive forms

Next we consider the family *M* of all strictly monotonic continuous forms. In other words, preferences should only satisfy transitivity, completeness, continuity, and monotonicity (Assumption 1.1). We saw before that sequential consistency imposes some specific restrictions when *M* is the RDU family or the betweenness family. Sequential consistency does not impose any further restrictions, other than those of separability as immediately imposed by the folding back assumption (1.1), for the family of monotonic continuous forms. So any general monotonic continuous forms V_E, V_{E^c}, V can be used in folding back (1.1), and in the single-stage evaluation (1.2) any functional W of the form $V(V_E(x_1,x_2), V_{E^c}(x_3,x_4))$ can be used. This is immediate and requires no proof.

Remark 6.1 Suppose that Assumption 1.1 holds and that folding back (Formula 1.1) holds with respect to the family *M* of monotonic continuous forms. Then sequential consistency is satisfied. □

7. Arbitrary finite numbers of events and stages

For simplicity of the exposition we have presented the main results for two stages and two-branch nodes. These results readily extend to more than two stages and more than two branches per node. A case with many branches is depicted in Figure 3a. There the state space is a set of the form $S \times T$ where S contains any finite number $n \geq 2$ of events and T contains any finite number $m \geq 2$ of events. We shall deal with a slightly more general

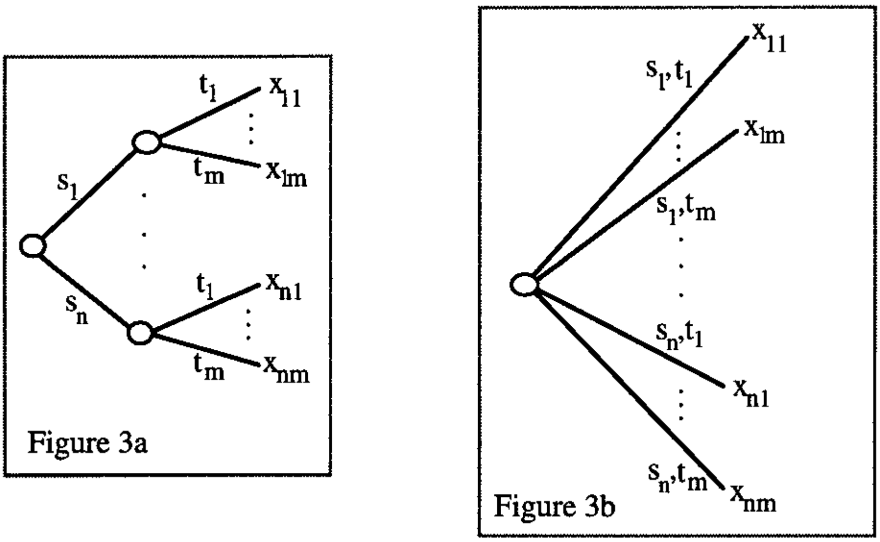

Figure 3.

case, illustrated in Figure 4a. There are n first-stage events $B_1,\cdots,B_n$, there are m_1 second-stage events $B_{11},\cdots,B_{1m_1},\cdots$, and there are m_n second-stage events $B_{n1},\cdots,B_{nm_n}$. In the folding back approach a representation of the form

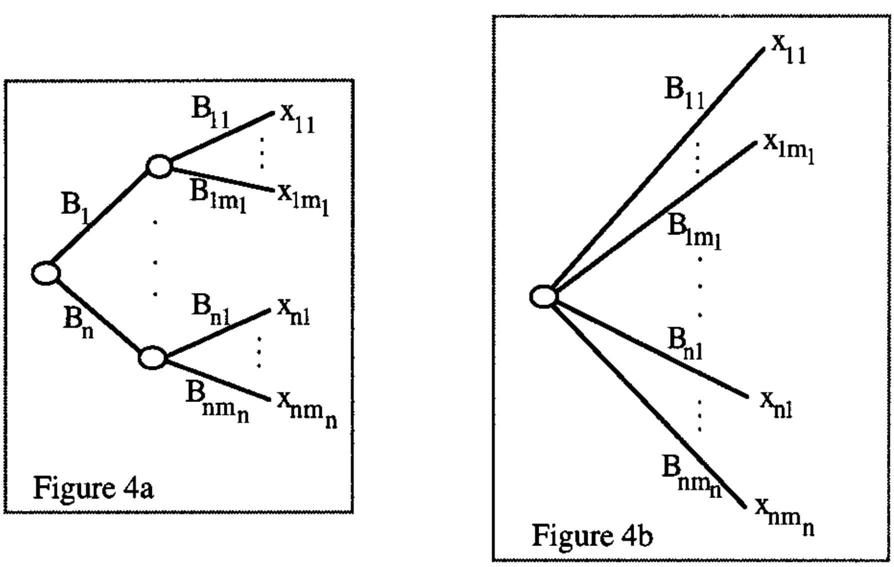

Figure 4.

$$V(V_1(x_{11},\cdots,x_{1m_1}),\cdots,V_n(x_{n1},\cdots,x_{nm_n})) \tag{7.1}$$

is imposed, where V and each V_j are taken from the family *M*. Note that this only imposes a weak form of separability. Each B_j is separable, but separability of other subsets of $\{B_1,\cdots,B_n\}$, such as $\{B_1,B_2\}$, is not imposed. We note that for the rank-dependent family it will follow, as a consequence of the other assumptions, that each subset of $\{B_1,\cdots,B_n\}$ is separable nonetheless. In the single-stage evaluation with respect to *M*, a representation of the form $W(x_{11},\cdots,x_{1m_1},\cdots,x_{n1},\cdots,x_{nm_n})$ is imposed, where W is an element of *M*. Again, sequential consistency requires that, if folding back holds with respect to *M*, then the single-stage evaluation is also with respect to *M*.

The extension to multiple events is straightforward for the general transitive forms (Theorem 5.1). The other models are described in more detail in the following theorem.

Theorem 7.1. Suppose that Assumption 1.1 holds and that folding back (Formula 7.1) and sequential consistency hold with respect to a family *M*, and that $n \geq 2$ and $m_j \geq 2$ for each j.

(i) If *M* is the multiple priors family, then sequential consistency is satisfied if all forms V and V_j use the same utility function and W uses the reduced family of probability measures (treating probabilities at different nodes as independent and multiplying them) and the same utility function.
(ii) If *M* is the rank-dependent family, then V is an EU form and all CEU forms V_j use the same utilities as V.
(iii) If *M* is the betweenness family, then each V_j is an EU form, the utilities U^k for $\succcurlyeq$ are independent of k ($U^k = U$ for all k), and each V_j uses the same utility U. V can be any form satisfying betweenness and using the same utility U in all EU descriptions of indifference classes.
(iv) If *M* is the weighted utility family, then V and all V_j's are EU forms.

□

Appendices D and F comment on the case of infinitely many branches. We next turn to the case of multi-stage trees. For multiple priors, the extension is straightforward again. The set of probability measures in the single-stage evaluation is now the reduced family of the sets in all intermediate chance nodes in the decision tree, i.e., the reduced forms of all combinations of probability measures at intermediate chance nodes. Also the general transitive forms (Theorem 5.1) readily extend to multiple stages. The other cases are described in detail in Appendix E. For the rank-dependent family, the conclusion is that only in the last stage can one deviate from EU, for the betweenness family the deviation from EU can only take place in the first stage, and for weighted utility no deviation from EU is possible.

8. Discussions and related literature

Consistency in dynamic choice situations has long been of interest in economics and decision analysis. With the advent of nonEU models, a renewed interest in dynamic consistency has arisen, as the application of nonEU models to dynamic settings poses new challenges. We have shown that a nonEU model (multiple priors) can be applied in dynamic settings much the same way as EU is applied. Some other nonEU models (rank-dependent and betweenness) can only be applied in a restrictive way. We now compare our results to those in the literature.

For the purpose of the following discussion we define invariance. Compare Figure 2a with Figure 5. In Figure 5, the order of events has been reversed, but corresponding strategies can be defined in the sense of assigning the same outcomes to the same states.[1] Invariance requires an identical evaluation of corresponding strategies in the two figures. When folding back is assumed both in Figure 2 and in Figure 5, together with invariance, then complete separability of preference over strategies is implied, which is the main condition leading to expected utility (Hammond 1988, Sarin & Wakker 1994). For the context of risk, invariance and weaker versions thereof are studied by Segal (1993). He uses the term "general order indifference" for invariance and points out that this condition is somewhat weaker than the reduction of compound lotteries axiom.

There is an abundance of research papers that delineate the assumptions under which independence or EU holds in dynamic choice situations (Hammond 1988, Chew and Epstein 1989, Karni and Schmeidler 1991, Sarin 1992, La Valle 1992). In our framework the assumptions that imply independence boil down to a universal commitment to dynamic consistency, consequentialism, and invariance. Therefore nonEU models must eschew a universal commitment to at least one of these three assumptions. Our results apply when any of these three assumptions is relaxed.

We first turn to the case where consequentialism and dynamic consistency hold universally (thus folding back can be employed throughout) but preferences are influenced by the order and timing of events and invariance is violated. For this case we have shown that

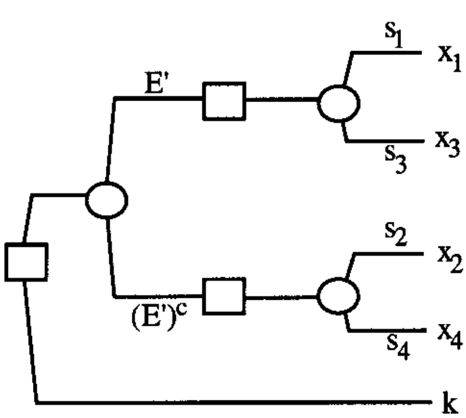

Figure 5. E' = $\{s_1,s_3\}$; $(E')^c = \{s_2,s_4\}$

dynamic choice can be evaluated using folding back and sequential consistency much the same way as in EU if one uses the multiple priors model. We agree with Kreps (1990, p. 114) that “The timing of resolution of uncertainty can be important for “psychological” reasons as well, even if the consumer has no particular use for the information in the sense that it could change some decisions he must make.” Violations of invariance have been studied by Chew and Epstein (1989), Grant, Kajii, and Polak (1998), Kreps (1979), Luce and Narens (1985), Luce and Fishburn (1991), Luce and von Winterfeldt (1994), and Segal (1987, 1990). Grant, Kajii, and Polak (1997) obtained a result that is remarkably similarly to ours. They explain violation of invariance by means of an intrinsic value of information and their conditions, though quite different than ours, also imply that for rank-dependent and betweenness models, deviations from EU can only occur in one stage. Empirical support for folding back has been given by Kahneman and Taversky (1979), Starmer and Sugden (1991), Bernasconi (1994), and others.

Second, our results extend the work of Machina (1989) and McClennen (1990) (see also Strotz 1956, p. 165, “precommit”) by showing that nonEU maximizers can achieve dynamic consistency while employing folding back or backward induction in evaluating alternative plans in some scenarios, but not necessarily in all. This non-universal folding back is analogous to preferences for commodity bundles that are not universally (“completely”) separable. Yet, some subsets of commodities can be separable. For example, commodities $\{x_1,x_2\}$ and $\{x_3,x_4\}$ may be separable whereas $\{x_1,x_3\}$ and $\{x_2,x_4\}$ are not. Folding back is analogous to Strotz's (1957) scheme of two-stage budget allocations, where in the first stage the budget is allocated between the two commodity groups that are separable, i.e., $\{x_1,x_2\}$ and $\{x_3,x_4\}$. In the second stage, inter-group budget allocations are determined; see also Gorman (1968). For nonseparable commodity groups $\{x_1,x_3\}$ and $\{x_2,x_4\}$, this two-stage budget allocation procedure (folding back) cannot be used because the allocation to a group depends on the distribution of budget within the other group.

Third, interest has also been expressed in the literature for violations of dynamic consistency. For instance, in the approach of Karni and Safra (1990), consequentialism and invariance are assumed to hold universally but dynamic consistency is relaxed. Jaffray (1994) presents an updating method and shows that dynamic consistency must be abandoned to implement the method. Strotz (1956) was perhaps the first modern economist to discuss a violation of dynamic consistency: “our present answer is that the optimal plan of the present moment is generally one which will not be obeyed, or that the individual's behavior will be inconsistent with his optimal plan” (p. 166). Allais’ (1953) arguments for deviating from EU are also based on a violation of dynamic consistency. He argues that choices made “ex ante,” i.e., prior to the revelation of uncertain information, need not conform to choices made “ex post,” i.e., posterior to the revelation of uncertain information. Similar to the example where universal consequentialism was relaxed, relaxation of dynamic consistency does not imply that dynamic consistency is inappropriate in all decision trees. In those trees where it is appropriate, our results can be applied.

To summarize, our results are useful in two ways. The first case occurs when a natural order decision tree representing a decision problem is to be evaluated using folding back, as in the product development example in Figure 1. In this case, the multiple priors model can achieve sequential consistency while permitting folding back. Second, our results show that one need not abandon consequentialism or dynamic consistency universally to be able to use a nonEU model.

Dynamic choice problems have mostly been studied in the context of decision under risk, where uncertainties are described by probabilities rather than by events. Let us emphasize that sequential consistency does not assume or imply reduction of compound lotteries. Risk is a special case of uncertainty (Wakker 1990), so our results are applicable to risk as well (for infinitely many outcomes, see Remark D.1 in Appendix D). Our results imply again that, for RDU, deviations from EU are only possible in final stages and for betweenness deviations are only possible in first stages.

The main point of our approach is that a decision maker may not universally commit to any of the principles of dynamic consistency, consequentialism, or invariance, but may violate each of them in certain specific situations. The flexibility of exploiting separability whenever possible but permitting nonseparability of preferences in other cases is the strength of our approach.

We now discuss related literature. Luce and Narens (1985) and Luce and Fishburn (1991) derive several nonEU models by relaxing various "accounting postulates." Accounting postulates mean that the sequence and the timing of events do not matter to a decision maker, and thus a relaxation of accounting postulates means a relaxation of invariance (Luce and von Winterfeldt 1994, p. 267). Luce's approach, however, involves other primitives such as a joint receipt operation and independent repetitions of events. "Recursive utility," surveyed in Epstein (1992), is similar. Again, the sequencing of events is relevant, i.e. the timing of the resolution of uncertainty may matter to the decision maker. Recursive utility applies to decision under risk and means that folding back is applied using the same forms $V = V_E = V_{E^c}$ in both the first and the second stage.

Epstein and Le Breton (1993) have derived probabilistic sophistication from dynamic consistency and some other assumptions, generalizing an earlier result of Machina and Schmeidler (1992); see also Grant (1995). Probabilistic sophistication means that beliefs can be quantified by additive probabilities. Epstein and Le Breton use all of Savage's (1954) axioms except the sure-thing principle but drop consequentialism. Therefore, in their approach folding back cannot be applied. Another difference between their approach and ours is that they impose their conditions universally, i.e., for all events, whereas we impose our conditions only for fixed first- and second-stage events. For those events, we preserve both dynamic consistency and consequentialism. In our results, it is possible to use nonadditive probabilities (see Theorem 3.1) that do not conform to probabilistic sophistication while still satisfying dynamic consistency. Analogous to our sequential consistency, Epstein and Le Breton also use a meta-principle to the effect that the same general family of models should be applicable at several decision points. Lo (1996) and Eichberger and Grant (1997) demonstrated that quadratic utility can satisfy the meta-principle of Epstein and Le Breton. Again, their forms do not satisfy consequentialism or sequential consistency and cannot be used in folding back.

A meta-principle can also be recognized in the conjugateness condition for prior probabilities in Bayesian statistics. Conjugateness of a family of probability distributions (e.g., the beta family) means that if the prior probability belongs to that family, then so does the posterior after updating (Winkler 1972).

Backward induction and other rationality principles for dynamic choice situations have been studied extensively in game theory (van Damme 1983, Kohlberg and Mertens 1986, Elmes and Reny 1994). Sequential consistency can be compared to subgame perfectness from game theory, introduced by Selten (1965). A commonly accepted assumption for extensive games was that all essential information is contained in the normal form of the game and that the normal-form solution should be a member of the family M of Nash equilibria. Selten argued that, given universal commitment to Nash equilibrium, it is equally appropriate for a solution that the prescriptions in each decision node in the extensive form, and the subgame emanating therefrom, be members of M.

9. Conclusion

Several papers in the literature have shown that when one makes a universal commitment to consequentialism, dynamic consistency, and invariance, then expected utility (independence) is implied. Those who are sympathetic to nonexpected utility find that implication overly restrictive because in their view violations of expected utility occur precisely because people do not universally satisfy all of the mentioned principles simultaneously. In order to give nonexpected utility models a fair shake, we do not impose these principles universally. Instead, we assume that the evaluation of a decision tree be carried out by folding back and that the same class of nonexpected utility models be used throughout the tree, and we do not impose restrictions on trees with other (orderings of) events. We have shown that the multiple priors model can preserve dynamic consistency, consequentialism, and sequential consistency for a general multiple stage tree. For a given tree, the multiple priors model can be applied much the same way as the expected utility model. Rank-dependent and betweenness models cannot be applied in that way. To the degree that the implications of sequential consistency for nonexpected utility models are considered undesirable, the condition can be considered an argument in favor of expected utility.

The implication of sequential consistency can be studied for alternative nonexpected utility models such as cumulative prospect theory (Tversky and Kahneman 1992), the α-Hurwicz (1951) criterion where a convex combination of minimal and maximal expected utility is taken, Jaffray's (1989) belief function theory, Gul's (1991) disappointment theory, or any other nonexpected utility model.

Sequential consistency can be used to develop new nonexpected utility models. For example, one can define the smallest sequentially consistent family that contains *all* Choquet expected utility forms for the folding back approach, or all betweenness forms. Thus, for each family of decision models, sequential consistency poses two research questions. First, what restrictions does sequential consistency impose on the family, second, what is the smallest family of models containing the original family that satisfies sequential consistency? This opens new research questions for studying dynamic choice under nonexpected utility.

Appendix A. Proof of Theorem 2.1

Assume that in the folding back procedure in Fig. 2a, probability measure P_4' minimizes expected utility at node 4, P_6' at node 6, and then P_2' at node 2. Next consider the minimization at node 7. Every probability measure P_7 from $\mathcal{P}_7$ results from choosing a probability measure P_2 from $\mathcal{P}_2$, P_4 from $\mathcal{P}_4$, and P_6 from $\mathcal{P}_6$, and multiplying probabilities. The resulting EU at node 7 is a convex combination of the EU values at nodes 4 and 6, hence is monotonic in the EU values at nodes 4 and 6. Whatever fixed probability measure P_2 is chosen at node 2, the minimal EU results if at nodes 4 and 6 the probability measures P_4' and P_6' are chosen. Given that, the probability measure P_7 at node 7 that minimizes expected utility is obtained by taking P_2' at node 2 and using P_2', P_4' and P_6' to obtain P_7'. □

Appendix B. Proofs and comments for Theorem 3.1

Proof of Theorem 3.1. We assume folding back and sequential consistency, and derive statement (iii) in the theorem. The other implications then are straightforward. It is also easily seen that Statement (iii) actually implies folding back.

Assume that U and ν are the utility and capacity used by the CEU form W in (1.2).

Remark B.1. The only implication of (1.1) used in the proof is that $E = \{s_1,s_2\}$ is separable for preferences over strategies.

Separability of $\{s_3,s_4\}$, also implied by (1.1), is not used in the proof. The main part of the proof will consist of deriving:

$$\nu(B) = \nu(B \cup C) - \nu(C) \tag{B.1}$$

for either $B \subset \{s_1,s_2\}$ and $C \subset \{s_3,s_4\}$ or $B \subset \{s_3,s_4\}$ and $C \subset \{s_1,s_2\}$. This will be done at the end. It shows, loosely speaking, that subsets of $\{s_1,s_2\}$ and of $\{s_3,s_4\}$ do not affect each other's decision weights.

Suppose now that (B.1) has been proved. (B.1) implies the equality $\nu(s_3,s_4) = \nu(s_1,\cdots,s_4) - \nu(s_1,s_2)$, which implies that $\nu(E) + \nu(E^c) = 1$, i.e., ν is additive with respect to E,E^c. Define "probabilities" $P(E) = \nu(s_1,s_2)$ and $P(E^c) = \nu(s_3,s_4)$. Define the capacity ν_E over $\{s_1,s_2\}$ as the normalized restriction of ν, i.e., $\nu_E = \nu/P(E)$. Similarly, define ν_{E^c} on $\{s_3,s_4\}$ by $\nu_{E^c} = \nu/P(E^c)$. For a strategy x, x_E denotes the *conditional act* assigning x_1 to s_1 and x_2 to s_2, and x_{E^c} denotes the *conditional act* assigning x_3 to s_3 and x_4 to s_4.

Letting U be the utility used by CEU in the single-stage approach, define $CEU_E(x_E)$ as the Choquet expected utility of the conditional act x_E over $\{s_1,s_2\}$ with respect to the capacity ν_E and the utility U; $CEU_{E^c}(x_{E^c})$ is defined similarly. It is straightforward from substitution and (B.1) that $CEU(x) = P(E)CEU_E(x_E) + P(E^c)CEU_{E^c}(x_{E^c})$, which is what Statement (iii) in the theorem requires.

We finally derive (B.1). We many assume that the range of U, which is an interval because of continuity, contains 0 in its interior. For simplicity of the notation, assume that outcomes are utilities, i.e., U is the identity. To avoid triviality, assume that B and C are nonempty. We first assume that B $\subset$ $\{s_1,s_2\}$ and C $\subset$ $\{s_3,s_4\}$ and treat the three cases $B = \{s_1\}$, $B = \{s_1,s_2\}$, and $B = \{s_2\}$ separately. Case 4 then considers the case where B $\subset$ $\{s_3,s_4\}$ and C $\subset$ $\{s_1,s_2\}$. We use a notation for strategies as displayed for instance in (B.2). There the strategy to the left of the $\sim$ indifference sign assigns outcome σ to event C, outcome to σ to s_1, outcome $-\tau$ to s_2, and outcome $-\tau$ to I.

Case 1. B = $\{s_1\}$. Write I = $\{s_3,s_4\}\backslash$C and choose $\sigma > 0 > -\tau$ such that

$$\begin{array}{cccc} C, & s_1, & s_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & \sigma & -\tau & -\tau \end{array} \sim \begin{array}{cccc} C, & s_1, & s_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & 0 & 0 & -\tau \end{array}. \tag{B.2}$$

Because of strong monotonicity and because the range of U contains an interval around 0, such outcomes can be found. By separability of $\{s_1,s_2\}$,

$$\begin{array}{cccc} s_1, & C, & s_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & 0 & -\tau & -\tau \end{array} \sim \begin{array}{cccc} s_1, & C, & s_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 0 & 0 & -\tau \end{array}. \tag{B.3}$$

In each indifference, we wrote the events in the order in which they appear in the CEU formula. The indifferences imply

$$\nu(\{s_1\} \cup C) - \nu(C) = \nu(s_1)\,, \tag{B.4}$$

because these are the two decision weights associated with outcome σ for the left strategies, which should outweigh the negative term in the CEU form provided by s_2. The latter is the same in both indifferences. (B.1) has been established for $B = \{s_1\}$.

In words, the indifferences (B.2) and (B.3) show that the decision weight of s_1 is not affected if C “crosses over” s_1 in the rank-ordering.

Case 2. B = $\{s_1,s_2\}$.

We choose $\sigma > 0 > -\tau$ as before, so have (B.2) and (B.3). Another application of separability of $\{s_1,s_2\}$ gives

$$\begin{array}{cccc} s_1, & s_2, & C, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & -\tau & -\tau & -\tau \end{array} \sim \begin{array}{cccc} s_1, & s_2, & C, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 0 & -\tau & -\tau \end{array}. \tag{B.5}$$

The indifferences (B.3) and (B.5) imply

$$\nu(\{s_1\} \cup C \cup \{s_2\}) - \nu(\{s_1\} \cup C) = \nu(s_1,s_2\}) - \nu(\{s_1\}) \tag{B.6}$$

because these are the two decision weights associated with outcome $-\tau$ given state s_2 for the left strategies, which should outweigh the positive term provided by s_1 that is the same in both indifferences. Adding up the equalities (B.4) and (B.6) gives $\nu(\{s_1\} \cup C \cup \{s_2\}) - \nu(C) = \nu(\{s_1,s_2\})$, i.e., (B.1) follows for $B = \{s_1,s_2\}$.

In words, the indifferences (B.3) and (B.5) showed that the decision weight of state s_2 is not affected if event C crosses over that event in the rank-ordering. We found in Case 1 that the decision weight of state s_1 is not affected if C crosses over that event in the rank-ordering. So C can cross over $\{s_1,s_2\}$ without affecting the decision weights of s_1 and s_2, thus of $\{s_1,s_2\}$.

Case 3. $B = \{s_2\}$. This case is similar to Case 1.

(B.1) has now been proved whenever $B \subset \{s_1,s_2\}$ and $C \subset \{s_3,s_4\}$. Case 4 can be proved similarly b invoking separability of $\{s_3,s_4\}$ instead of $\{s_1,s_2\}$. In order to prove Remark B.1, another short derivation is given.

Case 4. $B \subset \{s_3,s_4\}$ and $C \subset \{s_1,s_2\}$. (B.1) can be rewritten as $\nu(C) = \nu(B \cup C) - \nu(B)$ and then follows from the preceding cases.

The derivation of (B.1) essentially used the nonatomicity of $\{s_1,s_2\}$, which made nontrivial tradeoffs between s_1 and s_2 possible. In view of strong monotonicity that is assumed throughout this paper, *nonatomicity* of an event simply means that it is not singleton. □

The following result follows from Remark B.1 in the preceding proof.

Corollary B.2. If, under CEU and Assumption 1.1, a nonatomic event is separable, then so is its complement.

Proof. The general case follows similarly to the case of event E in the preceding proof for any two-fold partitioning $\{E_1,E_2\}$ of E with E_1 and E_2 nonempty so nonnull, instead of $\{s_1,s_2\}$. In the preceding proof, only separability and nonatomicity of E, and not of E^c, were used to derive Statement (iii). That in turn implies separability of E^c. □

An interesting implication of the corollary is that, under CEU, minimal separability assumptions, together with Gorman's (1968) result, imply separability of all events and then by Wakker (1996) subjective expected utility. The following corollary slightly strengthens Theorem 7.1.(ii) because it does not assume separability of A_n.

Corollary B.3. If, under CEU and Assumption 1.1, $(A_1,\ldots,A_n)$ is a partition of the state space, $n \geq 3$, and for all $j = 1, \ldots, n - 1$, event A_j is separable and nonatomic, then the capacity is additive on the algebra generated by $A_1,\ldots,A_n$.

Proof. Separability of A_j, $j = 1, \ldots, n - 1$ implies, by Corollary B.2, separability of all complements of these events. By Gorman (1968), all events measurable with respect to the partition are separable. By Wakker (1996, Theorem 3) the capacity is additive for those events. □

Appendix C. Proofs and comments for Theorem 4.1

Proof of Theorem 4.1 The implication (ii) ⇒ (i) is straightforward, hence we assume folding back and sequential consistency and derive (ii). Define $D^k = \{(x_1,x_2)$: $(x_1,x_2,x_3,x_4) \in W^{-1}(k)$ for some $x_3,x_4\}$, which can be rewritten as $\{(x_1,x_2)$: $W(x_1,x_2,M,M) \geq k \geq W(x_1,x_2,0,0)\}$.

Folding back implies separability of event E. We write $\succcurlyeq_E$ for the preference relation induced over pairs (x_1,x_2). Suppose, for a given indifference class $W^{-1}(k)$, that $p_1^k,\cdots,p_4^k$, U^k give the associated EU model for which $W^{-1}(k)$ is an indifference class. For k and $(x_1,x_2) \in D^k$, $(x_1,x_2) \sim_E (y_1,y_2)$ if and only if also $(y_1,y_2,x_3,x_4) \in W^{-1}(k)$, which holds if and only if $p_1^kU^k(x_1) + p_2^kU^k(x_2) = p_1^kU^k(y_1) + p_2^kU^k(y_2)$. We conclude for each k in the range of W and x_1,x_2,y_1,y_2 that, if $(x_1,\cdots,x_4) \in W^{-1}(k)$ for some x_3,x_4, then

$$(x_1,x_2) \sim_E (y_1,y_2) \Leftrightarrow p_1^kU^k(x_1) + p_2^kU^K(x_2) = p_1^kU^k(y_1) + p_2^kU^k(y_2). \tag{C.1}$$

The rest of the proof can, unfortunately, not be presented in a very accessible manner. The reason is that it is based on several advanced results from additive conjoint measurement theory that have been presented in different papers. Therefore we first present a special case that illustrates the gist of the proof. Consider the set K = {k: for each x_1,x_2 there exist x_3,x_4 such that $(x_1,\cdots,x_4) \in W^{-1}(k)\}$. For such k, (C.1) holds for all (x_1,x_2) and (y_1,y_2). Then, by monotonicity and continuity of $\succcurlyeq$, increasingness and continuity of U^k, and comparisons to certainty equivalents,

$$(x_1,x_2) \succcurlyeq_E (y_1,y_2) \Leftrightarrow p_1^kU^k(x_1) + p_2^kU^K(x_2) \geq p_1^kU^k(y_1) + p_2^kU^k(y_2). \tag{C.2}$$

Defining

$$q_1^k = \frac{p_1^k}{p_1^k + p_2^k}, \quad q_2^k = \frac{p_2^k}{p_1^k + p_2^k},$$

we see that q_1^k,q_2^k and U^k provide an EU representation for $\succcurlyeq_E$, for all $k \in K$. By standard uniqueness results (Wakker 1989, Observation IV.2.7'), this means that q_1^k and q_2^k are uniquely determined and that the U^k differ only by scale and location. Then we can, and will, choose such U^k identical. It means that we can suppress the superscript k in U^k and q_1^k, q_2^k for $k \in K$.

We have obtained an EU representation for $\succcurlyeq_E$ if K is nonempty. But we have also found:

There exists a utility U such that $U^k = U$ can be chosen for all $k \in K$. The proportion p_1^k/p_2^k is independent of $k \in K$.

We next turn to general k; actually, the above set K may be empty. (It can be seen that K is empty if and only if $(M,M,0,0) \prec (0,0,M,M)$.) Here we invoke advanced results from additive conjoint measurement. Take any (x_1,x_2) and set $k = W(x_1,x_2,z,z)$ for any $0 < z < M$. There exists an open neighborhood S of (x_1,x_2) such that for all $(y_1,y_2) \in S$ there exist (y_3,y_4) such that $(y_1,\cdots,y_4) \in W^{-1}(k)$. Then q_1^k, q_2^k, U^k provide an EU representation on S by a reasoning similar to the derivation of (C.1) and (C.2). In other words, for $\succcurlyeq_E$ there exist "local" representations. This implies the existence of a ("global") EU representation, as follows. First, by Chateauneuf and Wakker (1993, Corollary 2.3 −note that $[0,M]^2$ is an open subset for the topology restricted to $[0,M]^2$ − and Lemma C.1 and the text above that lemma), there exists a global additive representation for $\succcurlyeq_E$. Then, looking at the diagonal elements $(\alpha,\alpha) \in [0,M]^2$, their local EU representations, and standard uniqueness results, we find that the additive representation for $\succcurlyeq_E$ on $[0,M]^2$ is actually an EU representation. Let us display this:

There exist q_1,q_2 and U giving an EU representation for $\succcurlyeq_E$ on $[0,M]^2$.

We next turn to the EU representations on sets D^k through q_1^k, q_2^k, U^k. These EU representations represent the same preferences as the EU representation through q_1,q_2,U. The standard uniqueness results claim uniqueness of probabilities and uniqueness of utility up to scale and location on its domain, but these uniqueness results are only given on full product sets. The domain D^k is not a full product set. It is connected, as can be proved similarly to Chew, Epstein, and Wakker (1993, Lemma), and then it can be seen that the uniqueness result as above still holds (suggested by Wakker 1993, Section 2.2, discussion of "Second flaw"). We can conclude that $q_1 = q_1^k$ and $q_2 = q_2^k$ for all k, and $U^k = U$ can be taken by proper choice of scale and location for each k and U^k as defined through the domain D^k. Outside the domain D^k (so for real numbers not appearing as first or second coordinate of elements of D^k) the choice of U^k is only limited by strict increasingness and continuity, thus all U^k's can be chosen identical to U. We conclude:

$$U^k \textit{ can be chosen independently of } k \textit{ and } p_1^k/p_2^k \textit{ is independent of } k. \tag{C.3}$$

The proof so far has used no other implication of folding back than separability of E. That, together with betweenness of $\succcurlyeq$, apparently suffices to reduce the $\succcurlyeq_E$ representation to EU, giving (C.3). Similarly, using separability of E^c, we obtain an EU representation for $\succcurlyeq_{E^c}$. In the latter step, independence of utility from the superscript need no more be derived because it was already established in (C.3). (ii) follows. □

Appendix D. Proof and comments for Theorem 7.1

Proof of Theorem 7.1. For Statements (i), (iii), and (iv) in Theorem 7.1, the proofs of Theorems 2.1, 4.1, and 5.1 require no substantial modifications. For (ii), note that in the proof of Theorem 3.1 it was actually shown that, under CEU, separability of an nonatomic event E implies that the decision weights of subevents of E are not affected by outcomes outside of E. This result is applied here to $B_1,\cdots,B_n$ and then similarly implies Statement (ii). □

Let us consider Theorem 7.1 without the assumption $m_j \geq 2$ for all j. Then Statement (i), dealing with the multiple priors model, remains true; it requires no adaptation. The other statements, however, need no more be true. For example, if $m_j = 1$ for all j, then folding back (1.1) reduces to monotonicity and has no further implications, and Statements (ii), (iii), and (iv) are no more true. It can be seen that (ii) remains true if only one m_j is one but is no more valid if two or more m_j's are one. In the latter case one can, however, restructure the tree by grouping those m_j's together into one chance node from which the m_j branches emerge as subsequent branches. To this newly constructed tree, our result again applies. (The event belonging to the newly constructed chance node is separable indeed, as follows from Corollary B.3.)

Remark D.1. The extension to infinitely many branches is straightforward, by restrictions to simple acts (acts taking only finitely many values). The case where $S \times T = [0,1] \times [0,1]$, incorporating all two-stage probability distributions if S and T are endowed with the uniform probability distribution, is included. □

Appendix E. Multi-stage trees

We briefly describe how Theorem 7.1 can be extended to multi-stage trees. Assume a fixed non-terminal node B in a complex multi-stage tree, with n branches emerging from B, leading to "daughter"-nodes $B_1,\cdots,B_n$. Assume that these nodes B_j are not terminal nodes and from each B_j there emerge nodes $B_{j1}, \cdots, B_{jm_j}$. These nodes $B_{j1}, \cdots, B_{jm_j}$ are arbitrary. They may be terminal nodes, but also subtrees of any degree of complexity may emerge from them. In each of the following cases, we further assume:

(1) For each $1 \leq j \leq n$, $1 \leq i \leq m_j$, the outcomes associated with terminal nodes of all paths emerging from B_{j_i} are the same outcome, the generic notion for which is x_{j_i}.
(2) Outcomes associated with other terminal nodes, emerging from paths that do not pass through B, are fixed at some level μ.

Under these two assumptions, the resulting structure is isomorphic to the two-stage decision tree depicted in Figure 4a. With a slight abuse of terminology, we call the isomorphic two-stage structure a subtree. Dynamic consistency and folding back with respect to a family M in the multi-stage tree imply the same for the subtree and will be assumed henceforth (for the appropriate M in each case). The extension of the single-stage evaluation from the multi-stage tree to the subtree is considered for the following four cases.

(i) The extension of the multiple priors results is straightforward, still a reduced family of probability measures and the same utilities are used.

(ii) For the RDU family considered in Theorem 7.1(ii), one more restriction is imposed on the x_{j_i} outcomes in (1): we only consider the cases where $x_{j_i} \geq \mu$ for all j,i. Thus, the fixed outcome μ does not interfere with the rank-ordering of the other outcomes, comonotonicity in the multi-stage tree corresponds to comonotonicity in the subtree, and the single-stage evaluation in the subtree is also from the RDU family. Now we can apply Theorem 7.1(ii) to the subtree and conclude that, under the appropriate nontriviality conditions (at least two branches from each chance node; see also the remark on the restructuring in Appendix D), the decision maker maximizes EU at node B with the same utility U as at all daughter-nodes $B_1,\cdots,B_n$. The same can be concluded for every node B that is at least two steps away from terminal nodes.

(iii) For the betweenness family considered in (iii) of Theorem 7.1, we recall that betweenness means that each indifference class in the multi-stage tree is an indifference class of an EU form. Such an indifference class reduces to a corresponding EU indifference class in the subtree, implying a single-stage evaluation with respect to the betweenness family there also. Again, we can now apply Theorem 7.1(iii) under the appropriate nontriviality conditions and conclude that EU holds at all daughter-nodes B_j. This can similarly be derived for all non-beginning nodes B_k'. Therefore deviation from EU is only possible at the beginning node. Also the other results of Theorem 7.1(iii) readily extend.

(iv) For the weighted utility family in Theorem 7.1(iv), first note that this is a subfamily of the betweenness family. Therefore we can apply the results just established for betweenness and conclude, under appropriate nontriviality conditions, EU for all nodes except possibly the beginning node. The proof that also at the beginning node EU is maximized, is similar to the reasoning in Section 5: Weighted utility is an implicit utility model, where the EU models that describe indifference classes have varying utilities but all use the same probabilities. By the extension of Theorem 7.1(iii), however, such variations of utility are excluded and EU results also for the beginning node.

Appendix F. Solvability of events

In this section we assume a state space of the form $S \times T$ and a general outcome set $\mathscr{C}$, containing at least three nonindifferent outcomes. We adapt the results for the rank-dependent family to the case of infinitely many event-branches at the second-stage nodes of the decision tree, assuming Gilboa's (1987) solvability condition there. This adaptation makes it possible to discuss Schmeidler's (1989) two-stage approach. Given the infinite state space, null states are permitted hence we do not assume the strong monotonicity condition of Assumption 1.1 in this appendix. Given the possibly finite outcome set, we also drop the continuity condition of Assumption 1.1. The weak ordering condition of Assumption 1.1 is implied by the quantitative representation in folding back assumed in Theorem F.1.

In the folding back approach, we assume that the decision maker maximizes CEU with respect to the space T. More precisely, consider a preference relation $\succcurlyeq_T$ on the set of "conditional acts," i.e., functions from T to $\mathscr{C}$. Assume that a capacity ν_T is given on T and a utility U_T on $\mathscr{C}$, such that $\succcurlyeq_T$ is represented by CEU_T, the Choquet expected utility with respect to ν_T and U_T. In this section the richness assumption concerns the space T, instead of the outcome set $\mathscr{C}$. We assume that ν_T satisfies *solvability*, i.e., for all subsets $A \subset C$ of T and any $\nu(A) \leq b \leq \nu(C)$, there exists $A \subset B \subset C$ with $\nu(B) = b$. Gilboa (1987) used the term convex-ranged for this condition.

Now strategies are functions from $S \times T$ to $\mathscr{C}$. For simplicity, we assume that S is a finite set $\{s_1,\cdots,s_n\}$, for $n \geq 2$. We display the folding back assumption and the single-stage evaluation, for this setup:

(F.1) (folding back). There exists a Choquet expected utility representation CEU_T for $\succcurlyeq_T$ with a solvable capacity ν_T and a Choquet expected utility from V such that $\succcurlyeq$ is represented by

$$x \mapsto V(CEU_T(x_{s1}),\cdots,CEU_T(x_{s_n}))$$

where x_{s_j} denotes the *conditional act* assigning to each $t \in T$ the outcome $x(s_j,t)$.

(F.2) (single-stage evaluation). There exists a representation

$$x \mapsto CEU(x)$$

for $\succcurlyeq$, where CEU is a Choquet expected utility form.

Again, sequential consistency with respect to the rank-dependent family requires that (F.1) imply (F.2). Schmeidler (1989) considered the special case of folding back where the second-stage representation CEU_T is actually an expected utility form. The conclusion of the following theorem, however, is expected utility maximization at the first stage rather than at the second.

Theorem F.1. Suppose that folding back (F.1) holds and that there are at least three nonindifferent outcomes. Then sequential consistency with respect to the rank-dependent family holds if and only if there exist probabilities $p_1,\cdots,p_n$, and Choquet expected utility functional CEU and CEU_T, such that

$$CEU(x) = \sum_{j=1}^{n} p_j CEU_T(x_{s_j})$$

represents preferences.

Proof. For the "if" part, suppose that the representation in the theorem holds. For Assumption (F.2), we take the utility as in the CEU_T representation. For the capacity ν, let $E \subset S \times T$. Define $E_j \subset T$ such that $E \cap \{s_j \times T\} = s_j \times E_j$ for each j and set $\nu(E) = \sum_{j=1}^{n} p_j \nu_T(E_j)$. These definitions of U and ν give a CEU representation for (F.2).

Next assume folding back (F.1) and sequential consistency, which implies (F.2). We proceed to derive the representation in the theorem. Let ν and U be the capacity and utility, respectively, adopted by CEU in (F.2). We fix three nonindifferent outcomes $\alpha \prec \beta \prec \gamma$ and assume for now that $\mathscr{C} = \{\alpha,\beta,\gamma\}$. Restrict attention to the set of strategies that assign α to each state (s_j,t) for $j \geq 2$, and α, β, or γ to states (s_1,t). Both CEU_T and CEU give a Choquet expected utility representation for $\succcurlyeq_T$ (the latter with normalized capacity $\nu' = \nu/\nu(s_1 \times T)$). By Gilboa's (1987) uniqueness result, ν' and ν are identical capacities on T. Because this can be done for each s_j instead of s_1, we obtain the following solvability-like implication:

For each $s_j \times B$ and each $0 \leq \mu \leq \nu(s_j \times B)$ there exists a $B_1 \subset B$ with $\nu(s_j \times B_1) = \mu$. (F.3)

Another implication of Gilboa's (1987) uniqueness result is that we may assume that CEU and CEU_T use the same utilities for α,β,γ. (This holds for any triple $\alpha \prec \beta \prec \gamma$, from which it can be derived that CEU and CEU_T use the same utility on the entire set $\mathscr{C}$.)

The main part of the proof will consist of showing:

$$\nu(s_i \times B) = \nu((s_i \times B) \cup A) - \nu(A) \tag{F.4}$$

for all $B \subset T$ and $A \subset S\backslash\{s_i\} \times T$, and this will be done at the end. ((F.4) is similar to (B.1).) Let us now assume that (F.4) has been proved. This means for a subset $A \subset S \times T$ that, if we define $A_j := A \cap (s_j \times T)$, then $\nu(A) = [\nu(A_n \cup \cdots \cup A_1) - \nu(A_{n-1} \cup \cdots \cup A_1)] + [\nu(A_{n-1} \cup \cdots \cup A_1) - \nu(A_{n-2} \cup \cdots \cup A_1)] + \cdots + \nu(A_1) = \nu(A_n) + \cdots + \nu(A_1)$. Define, for each $s_j, p_j := \nu(s_j \times T)$ and define the capacity ν_j over T by $\nu_j(B) = \nu(s_j \times B)/p_j$ if $p_j > 0$ and ν_j is an arbitrary capacity if $p_j = 0$. $CEU_j(x_{s_j})$ denotes the Choquet expected utility of the conditional act x_{s_j} over T with respect to the capacity ν_j and the utility U. (F.4) means that the decision weights of subevents of $s_i \times T$ are independent of events outside of $s_i \times T$. From that and substitution we get $CEU(x) = \sum_{j=1}^{n} p_j CEU_j(x_{s_j})$.

As conditional on each s_j, the same preferences (represented by CEU_T) are induced over conditional acts, by standard uniqueness results all CEU_j must be the same functional as CEU_T for positive p_j's (remember they all adopt the same utility U) and can, and from now on will be the same functional as well for the p_j's that are zero. That functional is denoted as CEU_T. Note that it coincides with the functional CEU when restricted to strategies assigning the same conditional act to each s_j.

In the remainder of the proof, (F.4) will be derived, for $i = 1$. We may assume that values $\sigma > 0 > -\tau$ are contained in the range of U. For simplicity of the notation, let us assume that outcomes are utilities, i.e., that U is the identity. Define $I := (s_1 \times T \backslash B) \cup (\{s_2,\cdots,s_n\} \times T)\backslash A$. Hereafter, the outcome associated with I will be kept constant at a minimal level $-\tau$, so that the event I can be ignored for the choices and calculations. By (F.3), we can split $s_1 \times B$ up into $s_1 \times B_1$ and $s_1 \times B_2$, such that $\sigma\nu(s_1 \times B_1) = (\nu(s_1 \times B) - \nu(s_1 \times B_1))\tau$. This implies the indifference (with the notation of strategies explained above (B.2))

$$\begin{matrix} s_1 \times B_1, & s_1 \times B_2, & A, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & -\tau & -\tau & -\tau \end{matrix} \sim \begin{matrix} s_1 \times B_1, & s_1 \times B_2, & A, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 0 & -\tau & -\tau \end{matrix}. \tag{F.5}$$

The folding back representation (F.1) implies that replacements of the common outcome for event A do not change the indifference, so that

$$\begin{matrix} s_1 \times B_1, & A, & s_1 \times B_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & 0 & -\tau & -\tau \end{matrix} \sim \begin{matrix} s_1 \times B_1, & A, & s_1 \times B_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 0 & 0 & -\tau \end{matrix}. \tag{F.6}$$

and

$$\begin{matrix} A, & s_1 \times B_1, & s_1 \times B_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & \sigma & -\tau & -\tau \end{matrix} \sim \begin{matrix} A, & s_1 \times B_1, & s_1 \times B_2, & I \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \sigma & 0 & 0 & -\tau \end{matrix}. \tag{F.7}$$

In each indifference, we wrote the events in the order in which they appear in the CEU formula. The indifferences (F.6) and (F.7) imply, by the single-stage representation, that $\nu(s_1 \times B_1) = \nu((s_1 \times B_1) \cup A) - \nu(A)$, because these are the two decision weights associated with outcome σ for the left strategies, which should outweigh the negative term in the CEU form provided by $s_1 \times B_2$. The latter is the same in both indifferences. Similarly, the indifferences (F.5) and (F.6) imply $\nu(s_1 \times B) - \nu(s_1 \times B_1) = \nu((s_1 \times B) \cup A) - \nu((s_1 \times B_1) \cup A)$, because these are the two decision weights associated with outcome $-\tau$ given event $s_1 \times B_2$ for the left strategies, which should outweigh the positive term in the CEU form provided by $s_1 \times B_1$ that is the same in both indifferences. Adding up the two equalities yields $\nu(s_1 \times B) = \nu((s_1 \times B) \cup A) - \nu(A)$, i.e., (F.4) follows.

Repeating the derivation of (F.4) in words, (F.5) and (F.6) show that the decision weight of $s_1 \times B_2$ is not affected if A "crosses over" the event in the rank-ordering. (F.6) and (F.7) show that the decision weight of $s_1 \times B_1$ is not affected if A "crosses over" the event in the rank-ordering. Thus the joint decision weights of $s_1 \times B_1$ and $s_1 \times B_2$, i.e., the decision weight of $s_1 \times B$, is not affected if A "crosses over" the event in the rank-ordering. □

The following corollary was given, without proof, in Sarin and Wakker (1992). The expected utility model for probability distributions over prizes, assumed in the second stage of Schmeidler's (1989) model, can be considered a special case of a CEU model for random variables from the [0,1] interval, endowed with the Lebesgue measure, to the set of prizes. [0,1] is the state space in Gilboa's (1989) sense and the Lebesgue measure provides the solvable capacity.

Corollary F.2. The model of Schmeidler (1989) satisfies sequential consistency only if the capacity is additive and expected utility holds. □

Acknowledgments

The support for this research was provided in part by the Decision, Risk, and Management Science branch of the National Science Foundation.

Notes

1. In the context of our product development example, switching the order of events leads to a hypothetical and less natural tree. The decision maker may therefore want to adopt folding back in Figure 2 but not in Figure 5. This example illustrates the desirability of a flexible domain as considered in our analysis. The invariance condition is more natural if the temporal order of events is flexible.

References

Allais, M. (1953). "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Américaine," *Econometrica* 21, 503–546.

Ben-Porath, E., I. Gilboa & D. Schmeidler. (1997). "On the Measurement of Inequality under Uncertainty," *Journal of Economic Theory* 75, 194–204.

Bernasconi, M. (1994). "Nonlinear Preference and Two-Stage Lotteries: Theories and Evidence," *The Economic Journal* 104, 54–70.

Chateauneuf, A. and P.P. Wakker. (1993). "From Local to Global Additive Representation," *Journal of Mathematical Economics* 22, 523–545.

Chew, S.H. (1983). "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica* 51, 1065–1092.

Chew, S.H. (1989). "Axiomatic Utility Theories with the Betweenness Property," *Annals of Operations Research* 19, 273–298.

Chew, S.H. and L.G. Epstein. (1989). "The Structure of Preferences and Attitudes towards the Timing of the Resolution of Uncertainty," *International Economic Review* 30, 103–117.

Chew, S.H., L.G. Epstein, and P.P. Wakker. (1993). "A Unifying Approach to Axiomatic Non-Expected Utility Theories: Corrigenda," *Journal of Economic Theory* 59, 183–188.

Damme, E. van. (1983). "*Refinements of the Nash Equilibrium Concept.*" Springer, Berlin.

Dekel, E. (1986). "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory* 40, 304–318.

Eichberger, J. & S. Grant. (1997). "Dynamically Consistent Preferences with Quadratic Beliefs," *Journal of Risk and Uncertainty* 14, 189–207.

Eichberger, J. & D. Kelsey. (1996). "Uncertainty Aversion and Dynamic Consistency," *International Economic Review* 37, 625–640.

Elmes, S. and P.J. Reny. (1994). "On the Strategic Equivalence of Extensive Form Games," *Journal of Economic Theory* 62, 1–23.

Epstein, L.G. (1992). "Behavior under Risk: Recent Developments in Theory and Applications." In J.J. Laffont (Ed.), *Advances in Economic Theory* II, 1–63, Cambridge University Press, New York.

Epstein, L.G. and M. le Breton. (1993). "Dynamically Consistent Beliefs Must be Bayesian," *Journal of Economic Theory* 61, 1–22.

Epstein, L.G. and S.E. Zin. (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57, 937–969.

Fishburn, P.C. (1988). "*Nonlinear Preference and Utility Theory.*" Johns Hopkins University Press, Baltimore, MD.

Gilboa, I. (1987). “Expected Utility with Purely Subjective Non-Additive Probabilities,” *Journal of Mathematical Economics* 16, 65–88.

Gilboa, I. & D. Schmeidler. (1989). “Maxmin Expected Utility with a Non-Unique Prior,” *Journal of Mathematical Economics* 18, 141–153.

Gorman, W.M. (1968). “The Structure of Utility Functions,” *Review of Economic Studies* 35, 367–390.

Grant, S. (1995). “Subjective Probability without Eventwise Monotonicity: Or: How Machina’s Mom May Also Be Probabilistically Sophisticated,” *Econometrica* 63, 159–189.

Grant, S., Kajii, A., and Polak, B. (1998). “Intrinsic Preference for Information,” *Journal of Economic Theory*, forthcoming.

Grant, S., Kajii, A., and Polak, B. (1997). “Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models,” Working paper no 324, Dept. of Economics, Australian National University, Canberra, Australia.

Gul, F. (1991). “A Theory of Disappointment Aversion,” *Econometrica* 59, 667–686.

Hammond, P.J. (1988). “Consequentialist Foundations for Expected Utility,” *Theory and Decision* 25, 25–78.

Hazen, G.B. (1987). “Subjectively Weighted Linear Utility,” *Theory and Decision* 23, 261–282.

Hurwicz, L. (1951). “Optimality Criteria for Decision Making under Ignorance,” Cowles Commission Discussion Paper, Statistics, No. 370, mimeographed.

Jaffray, J.Y. (1989). “Linear Utility Theory for Belief Functions,” *Operations Research Letters* 8, 107–112.

Jaffray, J.Y. (1994). “Dynamic Decision Making with Belief Functions.” *In* R.R. Yager et al. (Eds), *Advances in the Dempster-Shafer Theory of Evidence*, 331–352, Wiley, New York.

Johnsen, T.H. & J.B. Donaldson. (1985). “The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans,” *Econometrica* 53, 1451–1458.

Kahneman, D. and A. Tversky. (1979). “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica* 47, 263–291.

Karni, E. and Z. Safra. (1990). “Behaviorally Consistent Optimal Stopping Rules,” *Journal of Economic Theory* 51, 391–402.

Karni, E. and D. Schmeidler. (1990). “Utility Theory with Uncertainty.” *In* W. Hildenbrand and H. Sonnenschein (Eds.), *Handbook of Mathematical Economics* 4, North-Holland, Amsterdam.

Klibanoff, P. (1995). “Dynamic Choice with Uncertainty Aversion,” Northwestern University.

Kohlberg, E. and J.F. Mertens. (1986). “On the Strategic Stability of Equilbria,” *Econometrica* 54, 1003–1037.

Kreps, D.M. (1979). “A Representation Theorem for Preference for Flexibility,” *Econometrica* 47, 565–577.

Kreps, D.M. (1990). “*A Course in Microeconomic Theory.*” Princeton University Press, Princeton, NJ.

Kreps, D.M. and E. Porteus. (1979). “Temporal Von Neumann-Morgenstern and Induced Preferences,” *Journal of Economic Theory* 20, 81–109.

LaValle, I.H. (1992). “Small Worlds and Sure Things: Consequentialism by the Back Door.” *In* W. Edwards (Ed.), *Utility Theories: Measurement and Applications*. Kluwer Academic Publishers, Dordrecht, 109–136.

Lo, K.C. (1996). “Weighted and Quadratic Models of Choice under Uncertainty,” *Economics Letters* 50, 381–386.

Luce, R.D. and P.C. Fishburn. (1991). “Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles,” *Journal of Risk and Uncertainty* 4, 29–59.

Luce, R.D. and L. Narens. (1985). “Classification of Concatenation Measurement Structures According to Scale Type,” *Journal of Mathematical Psychology* 29, 1–72.

Luce, R.D. and D. von Winterfeldt. (1994). “What common Ground Exists for Descriptive, Prescriptive and Normative Utility Theories,” *Management Science* 40, 263–279.

Machina, M.J. (1989). “Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty,” *Journal of Economic Literature* 27, 1622–1688.

Machina, M.J. and D. Schmeidler. (1992). “A More Robust Definition of Subjective Probability,” *Econometrica* 60, 745–780.

McClennen, E.F. (1990). “*Rationality and Dynamic Choice: Foundational Explorations.*” Cambridge University Press, Cambridge.

Quiggin, J. (1981). “Risk Perception and Risk Aversion among Australian Farmers,” *Australian Journal of Agricultural Economics* 25, 160–169.

Sarin, R.K. (1992). "What Now for Generalized Utility Theory." *In* W. Edwards (Ed.), *Utility Theories: Measurement and Applications*, 137–163, Kluwer Academic Publishers, Dordrecht.
Sarin, R.K. and P.P. Wakker. (1992). "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica* 60, 1255–1272.
Sarin, R.K. & P.P. Wakker. (1994). "Folding Back in Decision Tree Analysis," *Management Science* 40, 625–628.
Savage, L.J. (1954). "*The Foundations of Statistics.*" Wiley, New York. (Second edition 1972, Dover, New York.)
Schmeidler, D. (1989). "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587.
Segal, U. (1987). "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review* 28, 175–202.
Segal, U. (1990). "Two-Stage Lotteries without the Reduction Axiom," *Econometrica* 58, 349–377.
Selten, R. (1965). "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragetragheit," *Zeitschrift für die Gesamte Staatswissenschaft* 12, 301–324.
Starmer, C. and R. Sugden. (1991). "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation," *American Economic Review* 81, 971–978.
Strotz, R.H. (1956). "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies* 23, 165–180.
Strotz, R.H. (1957). "The Empirical Implications of a Utility Tree," *Econometrica* 25, 269–280.
Tversky, A. & D. Kahneman. (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.
Wakker, P.P. (1989). "*Additive Representations of Preferences*, A New Foundation of Decision Analysis." Kluwer Academic Publishers, Dordrecht.
Wakker, P.P. (1990). "Under Stochastic Dominance Choquet-Expected Utility and Anticipated Utility are Identical," *Theory and Decision* 29, 119–132.
Wakker, P.P. (1993). "Additive Representations on Rank-Ordered Sets II. The Topological Approach," *Journal of Mathematical Economics* 22, 1–26.
Wakker, P.P. (1996). "The Sure-Thing Principle and the Comonotonic Sure-Thing Principle: An Axiomatic Analysis," *Journal of Mathematical Economics* 25, 213–227.
Wald, A. (1950). "*Statistical Decision Functions.*" Wiley, New York.
Winkler, R.L. (1972). "*An Introduction to Bayesian Inference and Decision Theory.*" Holt, Rinehart and Winston, New York.